PNAS
www.pnas.org

Supplementary Information for

Ancillary Data for Co-citations In Context

James R. Bradley, Sitaram Devarakonda, Avon Davey, Dmitriy Korobskiy, Siyu Liu, Tandy Warnow, George Chacko

George Chacko.
E-mail: netelabs@nete.com

This PDF file includes:

Supplementary text
Tables S1 to S2

Supporting Information Text

This document contains the ancillary data that is referred to in the article *Co-citations In Context.*

Table S1. Hit Rates By Category

Data Set	Year	Highly Cited Min. Percentile	LNLC	LNHC	HNLC	HNHC
Immunology	1985	1%	0.0	0.8	0.6	**1.5**
Immunology	1985	2%	0.0	1.8	1.1	**3.0**
Immunology	1985	5%	1.0	3.5	3.2	**7.3**
Immunology	1985	10%	2.0	7.0	7.1	**14.1**
Immunology	1995	1%	0.0	**1.4**	0.5	**1.6**
Immunology	1995	2%	0.0	2.3	1.0	**3.0**
Immunology	1995	5%	0.0	4.4	3.2	**7.2**
Immunology	1995	10%	0.0	7.8	7.5	**13.3**
Immunology	2005	1%	0.0	0.6	0.6	**1.5**
Immunology	2005	2%	0.0	1.1	1.3	**2.8**
Immunology	2005	5%	0.0	3.0	3.5	**6.8**
Immunology	2005	10%	1.9	6.4	8.0	**12.8**
Metabolism	1985	1%	0.0	1.0	0.5	**1.6**
Metabolism	1985	2%	0.0	2.0	1.2	**3.0**
Metabolism	1985	5%	0.0	4.6	3.4	**7.2**
Metabolism	1985	10%	0.6	8.6	7.1	**13.9**
Metabolism	1995	1%	0.0	1.1	0.5	**1.5**
Metabolism	1995	2%	0.2	1.9	1.1	**3.0**
Metabolism	1995	5%	0.2	4.4	3.3	**7.1**
Metabolism	1995	10%	0.4	8.6	7.3	**13.7**
Metabolism	2005	1%	0.2	0.9	0.7	**1.3**
Metabolism	2005	2%	0.7	1.9	1.5	**2.7**
Metabolism	2005	5%	2.2	4.3	3.8	**6.6**
Metabolism	2005	10%	3.3	8.0	8.0	**12.6**
Applied Physics	1985	1%	0.0	0.4	**1.2**	**1.1**
Applied Physics	1985	2%	1.2	0.5	**2.4**	**2.3**
Applied Physics	1985	5%	3.5	2.3	5.4	**6.1**
Applied Physics	1985	10%	5.8	5.5	10.5	**12.5**
Applied Physics	1995	1%	0.3	0.5	**1.2**	**1.0**
Applied Physics	1995	2%	0.3	1.1	**2.4**	2.0
Applied Physics	1995	5%	1.3	2.7	**5.9**	5.2
Applied Physics	1995	10%	4.3	6.7	**12.2**	10.8
Applied Physics	2005	1%	0.5	0.2	**1.1**	**1.2**
Applied Physics	2005	2%	0.8	0.4	**2.1**	**2.4**
Applied Physics	2005	5%	1.0	1.4	5.3	**5.9**
Applied Physics	2005	10%	2.8	4.1	10.6	**11.6**
Web of Science	1985	1%	0.1	0.7	0.9	**2.2**
Web of Science	1985	2%	0.3	1.5	1.9	**4.2**
Web of Science	1985	5%	1.0	3.9	4.9	**10.0**
Web of Science	1985	10%	2.6	8.1	10.1	**18.4**
Web of Science	1995	1%	0.1	0.8	0.8	**1.9**
Web of Science	1995	2%	0.3	1.6	1.7	**3.7**
Web of Science	1995	5%	1.0	4.0	4.7	**9.1**
Web of Science	1995	10%	2.4	8.0	9.9	**17.3**
Web of Science	2005	1%	0.2	0.8	0.9	**1.7**
Web of Science	2005	2%	0.4	1.6	1.9	**3.3**
Web of Science	2005	5%	1.2	3.9	4.9	**8.2**
Web of Science	2005	10%	2.8	7.9	10.1	**15.5**

The hit rate is the percentage of publications in the referenced category that are in the top 1%, 2%, 5%, or 10% of papers according to citation count (see column 3) for novel articles defined as those with the 1st percentile z-score being negative. This table supplements the data reported in the SI document for novelty defined at the 10th percentile. The z-scores are computed using the local network. The category with the highest percentile is boldfaced (the second highest is also boldfaced if within 0.3% and greater than the overall percentage of articles considered to be hits).

Table S2. Explanatory Power of Novelty and Conventionality

Data Set	Year	Highly Cited Min. Percentile	Cumulative Probabilities: Conventionality Low	Conventionality High	Novelty Low	Novelty High
Immunology	1985	1%	1.000	**0.000**	0.874	0.414
Immunology	1985	2%	1.000	**0.000**	0.850	0.416
Immunology	1985	5%	1.000	**0.000**	0.997	0.267
Immunology	1985	10%	1.000	**0.000**	1.000	0.168
Immunology	1995	1%	1.000	**0.000**	0.240	0.583
Immunology	1995	2%	1.000	**0.000**	0.416	0.532
Immunology	1995	5%	1.000	**0.000**	0.939	0.379
Immunology	1995	10%	1.000	**0.000**	1.000	0.245
Immunology	2005	1%	1.000	**0.000**	0.961	0.391
Immunology	2005	2%	1.000	**0.000**	0.996	0.325
Immunology	2005	5%	1.000	**0.000**	1.000	0.260
Immunology	2005	10%	1.000	**0.000**	1.000	0.195
Metabolism	1985	1%	1.000	**0.000**	0.865	0.386
Metabolism	1985	2%	1.000	**0.000**	0.896	0.365
Metabolism	1985	5%	1.000	**0.000**	0.999	0.194
Metabolism	1985	10%	1.000	**0.000**	1.000	0.059
Metabolism	1995	1%	1.000	**0.000**	0.521	0.504
Metabolism	1995	2%	1.000	**0.000**	0.937	0.367
Metabolism	1995	5%	1.000	**0.000**	1.000	0.199
Metabolism	1995	10%	1.000	**0.000**	1.000	0.084
Metabolism	2005	1%	1.000	**0.000**	0.841	0.413
Metabolism	2005	2%	1.000	**0.000**	0.931	0.367
Metabolism	2005	5%	1.000	**0.000**	1.000	0.183
Metabolism	2005	10%	1.000	**0.000**	1.000	*0.042*
Applied Physics	1985	1%	0.177	0.860	0.999	0.105
Applied Physics	1985	2%	0.066	0.952	1.000	**0.023**
Applied Physics	1985	5%	0.200	0.818	1.000	**0.009**
Applied Physics	1985	10%	0.390	0.625	1.000	**0.001**
Applied Physics	1995	1%	0.062	0.955	0.999	0.113
Applied Physics	1995	2%	**0.018**	0.988	1.000	*0.036*
Applied Physics	1995	5%	**0.002**	0.999	1.000	**0.002**
Applied Physics	1995	10%	**0.000**	1.000	1.000	**0.000**
Applied Physics	2005	1%	0.319	0.706	1.000	**0.023**
Applied Physics	2005	2%	0.272	0.748	1.000	**0.004**
Applied Physics	2005	5%	0.117	0.897	1.000	**0.000**
Applied Physics	2005	10%	0.102	0.909	1.000	**0.000**
Web of Science	1985	1%	1.000	**0.000**	1.000	**0.000**
Web of Science	1985	2%	1.000	**0.000**	1.000	**0.000**
Web of Science	1985	5%	1.000	**0.000**	1.000	**0.000**
Web of Science	1985	10%	1.000	**0.000**	1.000	**0.000**
Web of Science	1995	1%	1.000	**0.000**	1.000	**0.000**
Web of Science	1995	2%	1.000	**0.000**	1.000	**0.000**
Web of Science	1995	5%	1.000	**0.000**	1.000	**0.000**
Web of Science	1995	10%	1.000	**0.000**	1.000	**0.000**
Web of Science	2005	1%	1.000	**0.000**	1.000	**0.000**
Web of Science	2005	2%	1.000	**0.000**	1.000	**0.000**
Web of Science	2005	5%	1.000	**0.000**	1.000	**0.000**
Web of Science	2005	10%	1.000	**0.000**	1.000	**0.000**

This table lists p-values in the form of cumulative right-hand tail probabilities for the observed number of hits in the Low Novelty, High Novelty, Low Conventionality, and High Conventionality categories under the sampling distribution generated by the null hypothesis of a random distribution of hit articles in proportion to the number of articles in each of the categories. A small p-value, therefore, indicates a number of hits that exceeds the expected number. Results that indicate statistically significant numbers of hits in excess of the expected number at the 0.05 level using a two-tailed test are highlighted in bold font, and those significant at the 0.10 level are italicized. These data are for the circumstances where novel citation patterns are defined by whether an article's 1st percentile z-score is negative, whereas the data for novelty being defined at the 10th percentile is included in the SI document. The z-scores are computed using the local network.